
Appendix. Source and Reliability of Estimates

SOURCE OF DATA

The estimates of school enrollment in 1983 are based on data obtained in October 1983 from the Current Population Survey (CPS) conducted by the Bureau of the Census. The CPS sample was initially selected from the 1970 census file and is updated continuously to reflect new construction where possible. The October sample was spread over 629 areas comprising 1,148 counties, independent cities and minor civil divisions with coverage in each of the 50 States and the District of Columbia. The sample is composed of approximately 61,500 occupied households that are eligible for interview. Of this number, about 2,500 occupied units were visited, but interviews were not obtained because the occupants were not found at home after repeated calls or were unavailable for some other reason. For a description of the CPS sample designs prior to 1980, see the detailed report for 1979 in this series.

The estimation procedure used for this survey involved the inflation of the weighted sample results to independent estimates of the civilian noninstitutional population of the United States by age, race and sex. These independent estimates are based on statistics from decennial censuses; statistics on births, deaths, immigration, and emigration; and statistics on the strength of the Armed Forces.

The independent population estimates used in this report to obtain data for 1983 are based on the 1980 decennial census.

Two sets of estimates for 1981 are shown in some of the tables in this report: one set results from using independent population estimates based on the more up-to-date 1980 decennial census and the other set results from using 1970 census based population estimates. The 1970 based estimates have been included to provide continuity in the time series with previous years. Comparing the 1980 based estimates with the 1970 based estimates provides a measure of the effect of changing to the 1980 based estimation procedure.

Data for 1972 to 1980 were obtained using independent population estimates based on the 1970 decennial census. Estimates for earlier years were based on earlier censuses.

RELIABILITY OF ESTIMATES

Since the estimates in this report are based on a sample, they may differ somewhat from the figures that would have

been obtained if a complete census had been taken using the same questionnaires, instructions, and enumerators. There are two types of errors possible in an estimate based on a sample survey— sampling and nonsampling. The standard errors provided for this report primarily indicate the magnitude of the sampling error. They also partially measure the effect of some nonsampling errors in response and enumeration, but do not measure any systematic biases in the data. The full extent of the nonsampling error is unknown. Consequently, particular care should be exercised in the interpretation of figures based on a relatively small number of cases or on small differences between estimates.

Use of school enrollment data for persons of Spanish origin. Methodological changes which occurred in 1980 resulted in relatively large increases in the estimated number and proportion of children 3 to 13 years old who are of Spanish origin. Consequently, when using school enrollment data for persons of Spanish origin, particular care should be exercised in comparing estimates for 1980 and later years of the total number of children enrolled in nursery school, kindergarten, and/or elementary school with estimates from earlier years. These changes do not affect school enrollment rates reported for persons of Spanish origin.

Sampling variability. The standard errors presented in tables A-1 and A-2 are primarily measures of sampling variability; that is, of the variations that occurred by chance because a sample rather than the entire population was surveyed. The sample estimate and its standard error enable one to construct interval estimates that include the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under identical conditions using the same sample design; and an estimate and its standard error were calculated from each sample, then:

1. Approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the average result of all possible samples.
2. Approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimate would include the average result of all possible samples.

Table A-1. Standard Errors of Estimated Numbers of Persons Enrolled in School, for the Total, Black, and Spanish-Origin Populations: October 1983

(Numbers in thousands)

Enrollment	Total		Black		Spanish origin	
	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error
Total enrolled.....	57,745	244	8,199	103	4,618	122
Nursery.....	2,350	72	326	28	108	16
Private.....	1,541	59	111	16	48	11
Kindergarten.....	3,361	85	476	33	335	28
Private.....	656	39	48	11	50	11
Elementary.....	27,198	132	4,153	93	2,548	75
Private.....	2,994	80	189	21	225	23
High school.....	14,010	153	2,143	63	1,104	62
College.....	10,825	137	1,102	48	523	37
Full time.....	7,711	119	806	41	335	28

Note: Controlled to 1980 census base.

Source: Estimates from table 1.

The average result of all possible samples may or may not be contained in any particular computed interval. However, for a particular sample one can say with specified confidence that the average result of all possible samples is included within the constructed interval.

All statements of comparison in the text have passed a hypothesis test at the 0.10 level of significance or better, and most have passed a hypothesis test at the 0.05 level of significance or better. This means that, for most differences cited in the text, the estimated difference between parameters is greater than twice the standard error of the difference. For the other difference mentioned, the estimated difference between parameters is between 1.6 and 2.0 times the standard error of the difference. When this is the case, the

statement of comparison will be qualified in some way; e.g., by use of the phrase "some evidence."

Note when using small estimates. Percent distributions are shown in this report only when the base of the percentage is greater than 75,000. Because of the large standard errors involved, there is little chance that percentages would reveal useful information when computed on a smaller base. Estimated numbers of persons are shown, however, even though the relative standard errors of these numbers are larger than those for the corresponding percentages. These smaller estimates are provided primarily to permit those combinations of the categories which serves each user's needs.

Table A-2. Standard Errors of Estimated Percentages of Persons 3 to 34 Years Old Enrolled in School for the Total, Black, and Spanish-Origin Population: October 1983

(Numbers in thousands)

Age	Total		Black		Spanish origin	
	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error
3 to 34 years.....	48.4	0.2	50.8	0.6	49.3	0.9
3 and 4 years.....	37.5	0.9	36.2	2.3	23.5	2.6
5 and 6 years.....	95.4	0.4	94.7	1.0	95.1	1.3
7 to 9 years.....	98.9	0.16	99.1	0.4	98.5	0.6
10 to 13 years.....	99.4	0.10	99.7	0.2	99.7	0.5
14 and 15 years.....	98.3	0.2	97.8	0.7	96.0	1.5
16 and 17 years.....	91.7	0.5	92.6	1.2	88.6	2.5
18 and 19 years.....	50.4	0.8	46.1	2.2	44.3	3.8
20 and 21 years.....	32.5	0.7	23.4	1.9	24.0	3.2
22 to 24 years.....	16.6	0.5	15.6	1.4	12.5	2.1
25 to 29 years.....	9.6	0.3	7.8	0.8	8.2	1.3
30 to 34 years.....	6.4	0.3	7.6	0.9	3.8	1.0

Note: Controlled to 1980 census base.

Source: Estimates from table 6.

Table A-3. Parameters to be Used for School Enrollment Characteristics for Direct Computation of Standard Errors

Characteristic	Parameters ¹	
	a	b
Persons enrolled in school, 3-34 years old:		
Total or White.....	-0.000017	2014
Black.....	-0.000117	2265
Spanish origin.....	-0.000028	3374
Persons enrolled in school, 14-34 years old:		
Total or White.....	-0.000025	2014
Black.....	.000179	2265
Spanish origin (level).....	² 0.001519	² 1856
Spanish origin (percent only)...	(X)	³ 3374
Children, 3-13 years old, enroll- ed in school:		
All races and Spanish origin...	-0.000063	2350

X Not applicable.

¹To obtain "a" and "b" parameters for 1965, multiply these parameters by 1.5.

²These "a" and "b" parameters are to be used to calculate standard errors of levels only for the October supplement. For Spanish origin in the March supplement, use a = 0.000901, b = 1101.

³This "b" parameter is to be used to calculate the standard error of percentages only. For the March supplement, use b = 2002.

Standard errors for data based on CPS. Since this is an advance report, standard errors are provided in tables A-1 and A-2, and standard error parameters are provided in table A-3 for estimated numbers of persons and estimated percentages for only certain characteristics which are considered the most important among the data in the report. A more complete

source and reliability statement for the 1983 data will be published with the forthcoming 1983 detailed report.

Standard errors of estimated numbers and estimated percentages. The approximate standard errors of estimated numbers and percentages can be computed directly with formulas (1) and (2) below, respectively. The formulas are:

$$\sigma_x = \sqrt{ax^2 + bx} \quad (1)$$

where "x" is the size of the estimate and "a" and "b" are the parameters associated with the characteristic; and

$$\sigma_{(x,p)} = \sqrt{\frac{b}{x} \cdot p(100-p)} \quad (2)$$

where "x" is the size of the subclass of the population which is the base of the percentage, "p" is the percentage (0 p 100), and "b" is the parameter associated with the characteristic.

Table A-3 provides the values of the "a" and "b" parameters that are used in formulas (1) and (2) to approximate standard errors of estimated numbers of persons and estimated percentages.

Standard error of a difference. For a difference between two sample estimates, the standard error is approximately equal to

$$\sigma_{(x-y)} \doteq \sqrt{\sigma_x^2 + \sigma_y^2} \quad (3)$$

Where σ_x and σ_y are the standard errors of the estimates x and y; the estimates can be numbers, percents, ratios, etc. This will represent the estimated standard error quite accurately for the difference between two estimates of the same characteristic in two different areas, or for the difference between two separate and uncorrelated characteristics in the same area. If, however, there is a high positive (negative) correlation between the two characteristics, the formula will overestimate (underestimate) the true standard error.